

Enabling spontaneous analogy through heuristic change

Article (Accepted Version)

Ormerod, Tom and MacGregor, James N (2017) Enabling spontaneous analogy through heuristic change. *Cognitive Psychology*, 99. pp. 1-16. ISSN 0010-0285

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/70537/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Running header: Spontaneous analogy

Enabling spontaneous analogy through heuristic change

Thomas C. Ormerod

University of Sussex, UK

James N. MacGregor

University of Victoria BC, Canada

Address for correspondence:

Thomas C. Ormerod

School of Psychology

University of Sussex

Falmer, BN1 9QH, UK

t.ormerod@sussex.ac.uk

Accepted for publication in *Cognitive Psychology*, September 2017

Abstract

Despite analogy playing a central role in theories of problem solving, learning and education, demonstrations of spontaneous analogical transfer are rare. Here, we present a theory of heuristic change for spontaneous analogical transfer, tested in four experiments that manipulated the experience of failure to solve a source problem prior to attempting a target problem. In Experiment 1, participants solved more source problems that contained an additional financial constraint designed to signal the inappropriateness of moves that maximized progress towards the goal. This constraint also led to higher rates of spontaneous analogical transfer to a superficially similar problem. Experiments 2 and 3 showed that the effects of this constraint extend to superficially and structurally different analogs. Experiment 4 generalized the finding to a non-analogous target problem that also benefitted from inhibiting maximizing moves. The results indicate that spontaneous transfer can arise through experience during the solution of a source problem that alters the heuristic chosen for solving both analogical and non-analogical target problems.

Keywords

Analogy; Transfer; Insight; Problem-solving; Heuristics

1. Introduction

Two influential papers on transfer in creative problem solving opened from apparently diametrically opposed positions. One, a seminal article on analogical transfer (Gick and Holyoak, 1980), began by asking “Where do new ideas come from? What psychological mechanisms underlie creative insight?” (p. 306) and went on to suggest transfer by analogy as the mechanism. The second paper argued that “...it is not transfer we want to achieve in the solution of important problems but freedom from transfer. The creative solution to an important problem may depend on freeing the problem solver from interference from old solutions...if we want to build creative problem solvers, should we teach people to transfer or teach them to avoid transfer?” (Detterman, 1993, p.2).

The transfer literature commonly focuses on the application of previously acquired skills and knowledge, and this active form of transfer is consistent with the Gick and Holyoak perspective. However, transfer may appear in another form, where what is transferred to a new situation is not the active application of a learned response but the capacity to inhibit a response or to suppress a quasi-automatic tendency. This latter perspective appears to be consistent with Detterman’s position, although in our interpretation it would not represent the avoiding of transfer but the transfer of avoiding previously dominant responses, or learning what not to do.

Insight is one area where learning what not to do may be particularly valuable. Insight is generally considered to involve problems in which an initial faulty or misleading representation precludes solving the problem until restructuring or representational change occurs. For example, Remote Associate Tasks (e.g., Bowman & Jung-Beeman, 2003) and Rebus problems (e.g., MacGregor & Cunningham, 2008)

often involve insight solutions, and solution may be inhibited when the problems are accompanied by words or phrases that prime a faulty representation (Smith, 1995; Smith & Blankenship, 1995). In such cases, the misleading representation is provided externally. In others, such as the nine-dot problem (e.g., Weisberg & Alba, 1991; MacGregor, Ormerod & Chronicle, 2001) and the six-coin problem (Ormerod, Chronicle & MacGregor, 2002), an initial misdirection may be caused by spontaneous cognitive tendencies that pre-exist within the problem solver.

1.1. Criterion of Satisfactory Progress Theory (CSPT)

Previously we proposed a theory of insight to explain why initial misrepresentations occur and how restructuring may be initiated via heuristic change (MacGregor, Ormerod & Chronicle, 2001; Ormerod, MacGregor, & Chronicle, 2002; Ormerod, MacGregor, Chronicle, Dewald & Chu, 2013). The theory proposed two heuristics that guide the selection of moves in problem solving, *move maximization* and *search minimization*, and a mechanism for evaluating candidate moves, *progress monitoring*. Move maximization is a tendency to select moves that maximize progress towards the perceived goal, while search minimization is a tendency to limit the problem space in order to reduce search for possibilities. Progress monitoring evaluates moves against a “criterion of satisfactory progress” derived from the problem description. A candidate maximizing move that is satisfactory with respect to the criterion will be executed.

To illustrate move maximization, consider the nine-dot problem, which consists of nine dots organized in a 3x3 array. Commonly, people attempt to draw through as many dots as possible with each line, such as using the first three lines to draw around three sides of the nine-dot figure. An attempt of this type meets the criterion of progress –each line cancels a minimum of the number of remaining dots divided by the number of remaining lines – until the final line is reached and solution

is seen to be impossible. However, up to this point of failure, the maximizing heuristic is so successful in meeting the criterion that subsequent attempts often show a similar response pattern, and the initial problem representation persists until a state of impasse is reached (MacGregor et al., 2001).

While the move maximization heuristic helps to explain why the nine-dot problem is so difficult, people do occasionally solve it (Weisberg & Alba, 1981; MacGregor et al, 2001). CSPT explains these occasional successes through *search minimization*, a second heuristic that operates to create and change the mental representation of a problem (Ormerod et al., 2013). Under search minimization, individuals limit the initial representation and subsequent expansion of the problem space to the minimum necessary to permit a search for moves that might meet the criterion. With the nine-dot problem, search minimization constrains the initial problem space to the dot array, which as described above allows for moves that meet the criterion but prevents a solution. Relaxing the search minimization heuristic allows a modification of the problem space, which for the nine-dot problem may include the space around the dot array, as well as other options that may result in illegal moves.

In applying CSPT to the nine-dot problem we further considered the effects of different levels of *mental lookahead* (MacGregor et al, 1999). A lookahead of one comprised the sampling and evaluating of one line, a lookahead of two, of two lines, and so on. However, in CPST, move selection and move evaluation involve different processes and it is theoretically possible for move selection to occur without evaluation. This happens when the problem solver employs insufficient cognitive resources to both mentally sample a move and then evaluate it. In the absence of any lookahead at all (“zero” lookahead), a maximizing move is considered to be selected

automatically without being evaluated against the criterion of progress (Ormerod et al, 2013). We found it necessary to make this distinction between sampling and evaluating moves when extending CSPT to the n -ball problem, described below, which forms the focus of empirical work reported here.

1.2. The n -Ball Problem

The most recent tests of CSPT used a variety of n -ball problems, where n can be between 7 and 9 balls depending on the variant used (Ormerod et al, 2013). The n -ball problem may be stated as follow: There are n apparently identical balls where one is imperceptibly heavier than the others. Using a balance-scale and limited to two weighs only, how can the heavier ball be identified?

For the 7-ball version, a solution is to weigh three balls against three (3v3). If the scale balances then the heavy ball is the one not weighed. If the scale is unbalanced then use the second available weigh to compare two of the three balls from the heavy side. If this weigh is balanced, then the heavy ball is the third ball. If it is unbalanced, then it is the ball on the heavy side of the scale. The solution to the 8- and 9-balls problems is almost identical: First, weigh 3v3. If the scale is unbalanced then use the second available weigh to compare two of the balls from the heavy side. If the scale is balanced, then use the second available weigh to compare two of the balls not previously weighed. If this second weigh is balanced, then the heavy ball is the third ball. If it is unbalanced, then it is the ball on the heavy side of the scale.

In the n -ball problem we defined a maximizing weigh as the one that maximized the number of balls in each pan of the balance-scale. We assumed that possible weighs are considered in order of decreasing maximization, starting with the maximizing weigh. For the 7-ball version, the first weigh to be considered (and

usually eliminated) is therefore $4v3$. Because all but one of n balls has to be eliminated after two weighs, we defined the criterion of satisfactory progress as $(n-1)/2$ balls eliminated after each weigh, on average. We considered a unit of lookahead to encompass the mental sampling and evaluation of one weighing, considering both a balanced and unbalanced outcome of the weigh. If a sampled move discovered under each level of lookahead meets the criterion, then it is selected and if not then the next weigh in order of decreasing maximization is selected (without evaluation). For example, with a lookahead of one, a weigh of $4v3$ is initially sampled and discarded, because mental evaluation shows that it results in an unbalanced outcome that eliminates no balls. It fails to meet the criterion of eliminating at least 3 balls, lookahead is exhausted, and the next weigh in order, $3v3$, is selected without evaluation.

The 8-ball and 9-ball versions of the problem are solved in a similar way to the 7-ball version. However, in both cases a $3v3$ weigh is not the maximizing move and a $4v4$ will tend to be preferred to a $3v3$ first weigh for problem solvers operating at lower levels of lookahead. For example, with an individual attempting the 9-ball problem at one-lookahead, the first weigh to be considered will be $5v4$. Evaluation shows that it eliminates no balls, and fails to meet the criterion. This exhausts lookahead and the next weigh, $4v4$, is selected without evaluation. This is an incorrect first weigh, whereas one-lookahead applied to the 7-ball resulted in selecting the correct first weigh. A number of predictions concerning the relative difficulty of problems and the frequencies of different first weighs were derived, and generally confirmed (Ormerod et al, 2013).

In the present article, we extend the CSPT perspective to examine transfer between different n -ball problems and between n -ball problems and both related

analogs and structurally unrelated problems. We present a theory that posits the withholding of move maximization that occurs under criterion failure as a potential mechanism for transfer. If withholding move maximization generalizes to different problems where the key to solution is to avoid maximizing moves, then we would expect to observe positive transfer.

1.3. Analogical transfer

Transfer is considered to occur when previously acquired knowledge is applied in a new learning or problem-solving situation. The fundamental nature of transfer provides, beyond psychology, tremendous practical potential to the fields of education and training. In the eyes of some recent reviewers, transfer remains, after more than a century, "...among the most challenging, contentious, and important issues for both psychology and education" (Day & Goldstone, 2012, p. 153).

Analogical transfer occurs when a learner or problem solver is influenced by commonalities between two disparate problems and applies solution relevant information from the source, or *base*, problem to the target problem (Catrambone, Craig & Nersessian, 2006; Day & Goldstone, 2011; Gick & Holyoak, 1980; Reeves and Weisberg, 1994). The degree of disparity between source and target may vary on a continuum from being virtually identical to being highly dissimilar. Transfer between problems towards the former pole of the continuum is characterized as *near* transfer, towards the latter, *far* transfer (Barnett & Ceci, 2002; Detterman, 1993). In addition to the number of common elements (Thorndike, 1906), the similarity of source and target may depend on the *depth* of the elements, which may vary from superficial to deep. Surface similarities involve concrete, domain-specific elements, and deep similarities, abstract, relational properties (Day and Goldstone, 2011). Major theoretical approaches differ in the extent to which they stress the roles of deep versus

surface similarities in analogical transfer. Theories such as structure-mapping (Gentner, 1983) and pragmatic schemas (Holyoak, 1985; Holyoak & Thagard, 1989) emphasize the formation of abstractions beyond surface similarities, while exemplar-based approaches (Medin & Ross, 1989; Reed, 1989; Ross, 1984) invoke the mediation of analogical transfer by both surface and deep problem similarities.

That similarity is a critical factor in analogical transfer is widely supported by the research evidence, and while structural similarity may play an important role, surface similarity appears frequently to determine the outcome of analogical transfer (Anderson, Farrell, & Sauers, 1984; Gentner, Ratterman, & Forbus, 1993; Holyoak & Koh, 1987; Ross, 1984). However, similarity does not always facilitate transfer and increasing similar content in base and target may not promote learning if it is irrelevant to the task (Gick & Holyoak, 1983; Quilici & Mayer, 1996; Reed, 1989). Also, transfer may be impeded if similar surface features are present but if they are cross-matched, playing different roles in the base and target situations (Gentner & Toupin, 1986; Ross, 1989). These potentially negative effects of surface similarity may depend on whether contextual factors promote a direct comparison of base and target (Lee, Betts, & Anderson, 2015). In addition, when the context encourages comparison, the effects may further depend on whether or not it causes over-elaboration of extraneous features in the mental representation of the problem (Bearman, Ormerod, Ball, & Deptula, 2011).

An additional distinction is made between *spontaneous* transfer, where the problem solver recognizes the relevance of the source information and applies it to the target problem without prompting, and *cued* transfer, where the problem solver is directly or indirectly informed of the source's relevance (Day and Goldstone, 2011; Didierjean and Nogry, 2004). Numerous studies have replicated this basic finding

(Bassok, 1990; Catrambone & Holyoak, 1989; Keane, 1985; Novik & Holyoak, 1991; Reed, Dempster & Ettinger, 1985), and generally, successful spontaneous transfer appears to be relatively rare in analogical problem solving (Day and Goldstone, 2011; Detterman, 1993; Barnet and Ceci, 2001). Difficulties in achieving spontaneous transfer have been observed apparently because participants fail to recognize the connection between the source and target problems without being cued to do so (Gick & Holyoak, 1980). Yet, there are demonstrations of spontaneous analogy in the literature (e.g., Clement, 1987; Blanchette & Dunbar, 2000; Bearman, Ball, & Ormerod, 2007). For example, Schunn & Dunbar (1996) found spontaneous analogical transfer in biological reasoning tasks despite participants being unaware consciously that they were using an analogical source to solve a target problem.

According to CSPT, if an individual appears to be making sufficient progress under their current criterion, then they may be reluctant to adopt strategies that require additional cognitive resources. Ormerod et al (2002) found that individuals did not make use of solution-relevant hints in solving insight problems until they had exhausted moves that appeared to make progress towards the problem goal. So, individuals may not always make use of analogs when simple strategies such as move maximization appear to be working. This problem is likely to be exacerbated where the surface features of source and target analogs differ (Gentner & Toupin, 1986; Holyoak & Koh, 1987; Ross, 1987), since an analogical strategy requires not only the mapping of problem and solution components between source and target, but also the retrieval of relevant analog problems. Thus, one way to unlock spontaneous analogical transfer may be to remove the incentive to continue with a move maximization heuristic and to seek more productive routes. The experiments reported below explore this hypothesis in relation to analogical transfer between variants of the

n -ball problem.

The question then arises as to how the limitations of move maximization might be discovered. Move maximization is frequently adaptive, since it offers the general problem-solving heuristic of hill-climbing (e.g., Russell & Norvig, 2004). According to the CSPT, insight problems like the n -ball and the 9-dot are difficult because (a) they encourage move maximization, yet (b) ensure that it will fail by the conditions they impose (4 lines only and no lifting of pen in 9-dot, 2 weighings only in 8-ball). The problem constraints are such that move maximizing will apparently be successful up until the goal is almost reached. Unless a participant has sufficient lookahead to anticipate failure, then going through all the steps leading up to failure may make it difficult to appreciate that maximization is getting in the way, and not some other aspect of the problem. Thus, the experience of failure to solve is unlikely on its own to be sufficient to dislodge a move maximization heuristic.

One approach to making participants aware of the limitations of move maximization is to provide additional problem constraints that induce criterion failure at low lookaheads. For example, MacGregor et al (2001, Experiments 4 and 5) gave participants a variant of the 9-dot problem in which the first line of a solution was already drawn, either as a horizontal line extending beyond the perimeter of the 9 dots, or as a diagonal line connecting top left and bottom right dots. Most theoretical accounts of insight problem-solving would predict an advantage for the horizontal version, since it embodies the ‘insight’ that lines must extend beyond the square of dots. In contrast, CSPT predicts an advantage for the diagonal line version, since solution attempts from that first line lead to criterion failure at any lookahead > 1 . In contrast a lookahead of 3 is required to anticipate criterion failure for the horizontal version. The predicted advantage for diagonal over horizontal versions was confirmed

by MacGregor et al's results. In essence, participants solving the diagonal version experienced the limitations of maximizing move selection before they ran out of lines to draw, while those solving the horizontal version did not.

With the *n*-ball, a signal of criterion failure at low lookahead levels may also help in dislodging a move maximizing heuristic. One mechanism for achieving this signal is to add an additional constraint that is irrelevant to the solution structure, in which a charge of \$1 is made for each ball weighed. If participants are given \$8, then this is sufficient to find the correct solution. However, if participants solving 8- and 9-ball variants select maximizing first moves, then their money is spent on the first move, preventing a second move attempt. If participants are given \$12, then the constraint does not impact on any weighings they may choose to make. Thus, the \$8 cost constraint shows in advance of failure that a 4v4 weigh will not work at all levels of lookahead > 1 . In effect, the limited cost constraint should expose participants to the limitations of maximizing ball weighs before they fail to solve the problem as a whole.

1.4. Transfer within and from *n*-ball problem variants

The first experiment reported below examined transfer between different versions of *n*-ball problems. Specifically, the experiment compared the effects of training on the 7-ball version against training on the 9-ball version on subsequent transfer to the 8-ball problem. Because *n*-ball problems tend to be highly similar to each other, both in superficial characteristics and in deeper, structural relations, transfer in the first experiment may be considered as lying towards the near pole of the near-far transfer continuum. The high level of both surface and relational similarity between them, the training and transfer problems in this case make them examples of literal similarity rather than of analogy (Gentner & Markman, 1997). As

such, their interconnections may be easier to recognize and to map, which might allow spontaneous analogical transfer to occur. However, as Gentner & Markman point out, literal similarities may have identical solutions that can be reproduced from solution memory without analogical mapping per se. In our experiments, we introduced a delay between source and target problems filled with another activity in order to clear solution memory. We predicted that, under these conditions, there would be competition between problem-solving strategies based on analogy and move maximization, in turn reducing spontaneous analogy rates.

The second and third experiments examined transfer between n -ball problems and structurally similar but superficially dissimilar problem analogs, and therefore involved transfer lying further towards the far pole of the transfer continuum. In constructing the transfer problems for these experiments, we sought to vary superficial differences between source and target by increasing categorical distance, to increase structural differences by varying the nature and number of objects, attributes and relations in source problems.

In Experiment 2, the analog problem involved testing nuclear rods to isolate one creating excess heat (see Appendix 1 for source and target problem statements). In this case, plutonium rods substituted for balls, a testing device substituted for the balance, and a difference in heat generation substituted for a difference in weight. Thus, although structurally identical, the source and target problems were superficially dissimilar. In Experiment 3, the analog problem involved isolating a sheepdog that had turned sheep killer, where sheepdogs substituted for balls, barns and leashes substituted for the balance, and the presence or absence of a dead sheep in a barn substituted for a difference in weight. The sheepdog scenario of Experiment 3 would appear to be a more superficially distant analog of n -ball problems than the

nuclear rods scenario of Experiment 2. Also, the objects ‘pans’ and ‘bins’ of Experiments 1 and 2 have attributes ‘on/off’ or ‘up/down’ that directly yield the test result. In contrast, the sheepdog problem has an object ‘barn’ that has the attribute ‘contains sheep’; the sheep are an additional problem object with the attribute ‘alive/dead’. As such, the sheepdog source problem is structurally similar to, but not identical with, the target problem, since the test result must be inferred from an additional level of object-attribute relations. Thus, transfer increased from relatively near in Experiment 1 to relatively far in Experiment 3.

Experiment 4 provided a critical test of the hypothesis that relaxing a progress maximizing heuristic may be a mechanism of transfer, by examining whether training on problems whose solutions required non-maximizing moves improved performance on a problem that similarly involved avoiding a maximizing move but otherwise bore neither superficial nor structural resemblances to the training problems. Experiment 4 used the 9-ball and sheepdog problems of Experiment 3 as sources and the Cheap Necklace Problem (CNP) as the transfer task (Chu, Dewald & Chronicle, 2007). The goal of the CNP is to combine four 3-link chains into a closed 12-link loop, where the cost of opening a link is 2 cents, the cost of closing a link is 3 cents and a total of 15 cents is available. (Chu et al., 2007). The solution requires unlinking all three links of one of the 3-link chains and closing them to connect the three remaining 3-link chains into a closed 12-link loop, for a total cost of 15 cents. A major difficulty in solving appears to be a tendency to directly connect together the original four 3-link chains. This maximizes progress towards the goal until the final step, when no more money is available to connect the 12-link chain into a closed loop (Chu et al.). In contrast, by unlinking a length of chain, the solution initially requires the antithesis of progress. In examining transfer from the n -ball problem (and an analog) to the unrelated CNP,

Experiment 4 therefore appears to exceed the far limit of the near/far continuum, and enter a domain of non-analogical transfer.

2. Experiment 1

In Experiment 1, participants received as source problems either 7-ball or 9-ball versions of the n -ball problem, and were subsequently tested on the 8-ball transfer problem. The maximizing weigh in the 7-ball problem is 3v3, which is also the correct first weigh, and solving the 8-ball target problem also involves an initial weigh of 3v3. If transfer can be mediated by the inhibition of a progress maximization strategy, and participants in the 9-ball condition learn not to select a maximizing move and transfer this knowledge to the 8-ball problem, then training on the 9-ball source problem will be more effective than training on the 7-ball.

Experiment 1 included a second independent variable, the presence or absence of a structurally irrelevant cost constraint whose presence was designed to signal the inappropriateness of a maximizing first weigh. A cost of \$1 was imposed for each ball weighed, with a total cost of \$8 allowed in one condition and \$12 in the other. The intention of this manipulation was to further suppress a tendency to maximize, and to test whether this suppression generalized to the transfer task. For the 9-ball condition, imposing an \$8 constraint should cause immediate criterion failure in participants operating with sufficient lookahead, since the maximizing 4v4 first weighing will exhaust the available resources, whereas the \$12 constraint should have no similar effect. This makes it more likely that, in the \$8 condition, the 3v3 weighing will be chosen (the correct first weighing), in turn resulting in more correct solutions. Therefore, the expectation is that, for the \$8 condition compared with the \$12, the incidence of maximizing (4v4) moves should be lower and the incidence of correct

solutions should be greater.

In the 7-ball condition, unlike the 9-ball, the \$8 constraint does not cause immediate failure, allowing the maximizing 3v3 weighing with \$2 to spare, leading to the prediction that the incidence of the 3v3 move and the ultimate solution rates should be the same for both the \$8 and \$12 constraint. The experimental predictions are therefore for a significant interaction between the independent variables, with the \$8 constraint, but not the \$12, having a facilitative effect on 9-ball but not 7-ball performance.

Ormerod et al (2013) found that a surprisingly large proportion of participants selected a seemingly irrational imbalanced first weighing (3v4 for the 7-ball problem, 4v5 for the 9-ball problem, etc.). In these cases, any possible inference about the presence of the heavier ball is masked by the uneven number of balls on the balance. Where an imbalanced weigh involved all available balls, we interpreted this as further evidence for a maximizing tendency with these problem types. To examine whether performance on the source problems was guided by a search for maximizing moves, we therefore examined the frequencies with which participants selected maximizing moves as their first weigh (3v3 and 3v4 for the 7-ball source, and 4v4 and 4v5 for the 9-ball source). Again, we predicted a reduction in maximizing moves under an \$8 constraint for the 9-ball source compared with \$12, but with no effect for the 7-ball source.

2.1. Method

Participants. One hundred and sixty-two first-year undergraduate students (26 male, 136 female, mean age 19.1 years, range 18-47 years) were tested as part of an introductory Psychology 101 class.

Materials. Each of the problems was presented in the same fashion with the

only difference being the wording as to the number of balls being weighed or money available for weighing. Each participant received an eight-page booklet containing, in the following order; general instructions, a participant information sheet, a source problem, a written solution to the source problem, a page instructing participants to set aside the booklet without referring to it again, a target problem, a written solution to the target problem, and debriefing information. The 7-ball and 9-ball source problems were described using the following instructions:

You have seven (nine) balls that look identical, yet one of them is slightly heavier than the others (but the difference is too small to detect just by picking them up). You have a balance-scale, and you can use it only twice (that is to say, you have two weighings at your disposal). Also, it costs \$1 to weigh each ball, and you have a maximum of \$8 (\$12) to spend. Your task is to find out which one is heavier, using only two weighings and without spending more than \$8 (\$12).

The form in which solutions were given was identical to each problem and included no cost information, as follows:

The solution to the problem you have just solved is as follows: Put three balls on each side of the scales. If the scales balance, then the unweighed ball is the heavier one. Otherwise, take two of the three balls that were on the downward side of the scale, and weigh one against the other. If the scales balance, then the unweighed ball is the heavier one. Otherwise, the ball on the side of the scales that dips downwards is the heavier one.

Design and Procedure. Participants were assigned to either 7-ball plus \$8, 7-ball plus \$12, 9-ball plus \$8 or 9-ball plus \$12 condition by alternating the order in which booklets were distributed across participants.

After completing the participant information sheet, the experimenter instructed participants to turn the page, and participants were allowed five minutes to attempt the source problem (7-ball or 9-ball). Participants wrote their solution attempts in the booklet, and were asked to record separately the first weighing immediately it occurred to them. After five minutes, the experimenter instructed participants to turn the page and to study the solution for two minutes, referring back to their own solution attempt if they wished. They were then told to put aside the booklet, and were reminded not to work on or discuss the training problem with anyone else until the experiment resumed. Participants then received a 45-minute lecture on literature search and reviews, after which they were instructed to pick up the booklet and attempt the target problem (the 8-ball problem) for five minutes. After five minutes, they were shown the solution and given debriefing instructions.

2.2. Results

Source solution proportions. For the 7-ball source, solutions rates were 28% (11/40) under the \$8 constraint and 26% (11/42) under \$12. For the 9-ball source, the corresponding rates were 55% (23/42) and 8% (3/38). A logistic regression using Source (7-, 8-, and 9-ball), Constraint (\$8, \$12) and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 162) = 11.12, p = .011$, with Source (Wald = 4.66, $p = .031$), Constraint (Wald = 4.07, $p = .044$), and the interaction between Source and Constraint (Wald = 5.09, $p = .024$) all significant predictors in the model. The \$8 constraint yielded more solutions (41%) than the \$12 (17%) overall, and the \$8 constraint resulted in higher solution rates than the \$12 with the 9-ball source, but not with the 7-ball.

Source first weighing frequencies. For the 7-ball source, the frequencies of criterion-satisfying first weighings (including imbalanced maximizing weighs such as

4v5 and 8v1 as well as 4v4 weighings) were 78% (31/40) under the \$8 constraint and 90% (38/42) under the \$12. For the 9-ball condition, the corresponding rates were 50% (21/42) and 84% (32/38). Although a logistic regression using Source, Constraint, and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 162) = 25.10, p < .001$, none of the predictors was significant in the model.

Target 8-ball solution proportions. For those previously given the 7-ball source, solution rates for the 8-ball target were 28% (11/40) for those who had experienced the \$8 constraint and 36% (15/42) for the \$12. For the 9-ball training condition, the corresponding transfer rates were 71% (30/42) and 47% (18/38). A logistic regression using Source, Constraint, and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 162) = 18.74, p < .001$, with Source (Wald = 4.36, $p = .037$) and the interaction between Source and Constraint (Wald = 7.30, $p = .007$) significant predictors in the model. Constraint (Wald = 3.80, $p = .051$) was not significant in the model.

Target 8-ball first weighing frequencies. For the 7-ball source, the frequencies of criterion-satisfying first weighings for the 8-ball transfer problem were 65% (26/40) under the \$8 constraint and 62% (26/42) under the \$12. For the 9-ball source, the corresponding rates were 29% (12/42) and 58% (22/38). A logistic regression using Source, Constraint, and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 162) = 14.36, p = .002$, with Source (Wald = 6.35, $p = .002$) and the interaction between Source and Constraint (Wald = 4.30, $p = .038$) significant predictors in the model. Again, Constraint (Wald = 3.45, $p = .063$) was not significant in the model.

2.3. Discussion

The first experiment compared transfer rates between different versions of n -ball problems, specifically contrasting the effects of training on the 7-ball versus 9-ball versions on subsequent transfer to the 8-ball problem. Because all three problems were highly similar, both in superficial characteristics and in deeper, structural relations, the first experiment represented an example of relatively near transfer. According to CSPT, maximization in the 7-ball problem will tend to elicit a first move that happens to be the correct first move for the 7-ball, 8-ball and 9-ball versions. In contrast, maximizing in the 9-ball version encourages an incorrect first move and a maximizing tendency has to be abandoned in order to solve the problem. The two training versions therefore set in opposition two different transfer possibilities. If what transfers is a tendency to reuse the correct first move, then the 7-ball version should be the more effective training problem. Conversely, if what transfers is a suppression of the tendency to maximize, then the 9-ball version should be more effective. Results supported the latter explanation, both in terms of the first moves made and solution rates.

An argument against this account is that predicting literal solution reuse for the 7-ball transfer condition may be a simplistic interpretation of analogical transfer. Another possibility is that participants who received 7-ball training and experienced the success of the 3v3 weighing may have had their tendency to maximize reinforced, resulting in poorer performance on the 8-ball transfer task. This interpretation is equally consistent with the data, although it differs from our preferred explanation in that differences on transfer would result from poorer transfer performance under 7-ball training rather than enhanced transfer performance under 9-ball training. Unfortunately, the experiment did not use a control condition that examined 8-ball

performance without training, which prevents testing between these possibilities. The second experiment, reported below, employed a control of this kind.

In addition, the present experiment introduced tight and lax cost constraints to further indicate failure of a maximizing first move in the 9-ball source problem under a tight cost constraint. This manipulation influenced performance on the source problem by resulting in fewer maximizing first moves and higher solution rates in the 9-ball, tight constraint, condition than in the other conditions. This influence carried over to performance in the transfer task, with higher solution rates for those trained on the 9-ball problem than the 7-ball, and with the highest rates occurring for those trained on the 9-ball problem under a tight cost constraint.

3. Experiment 2

The results of Experiment 1 provide evidence for spontaneous transfer between ball-weight problem analogs. However, the transfer problem in Experiment 1, the 8-ball variant, is both structurally and superficially similar to the 7-ball and 9-ball source problems. Thus, in Experiment 2, we introduced a structurally similar but superficially different target, the ‘Nuclear Rods’ problem, shown in Appendix 1. The introduction of this target problem enables a test of the mediating effects of solution experience on spontaneous analogical transfer. However, without knowing the solution rates for this problem in the absence of a preceding analogical source problem, the extent of analogical transfer cannot be assessed. Indeed, it is possible that the effects of preceding solution attempts with an analogous source problem might be negative. Therefore, data were collected from an additional group of participants, who solved only the nuclear rods problem, serving as a no-analog control.

In Experiment 2, the cost-constraint factor was expanded to three levels comprising \$8 (tight), \$12 (lax), and no cost constraints. As in Experiment 1, imposing an \$8 constraint in the 9-ball conditions should cause immediate criterion failure, since the preferred 4 balls against 4 first weighing will exhaust the available resources, whereas no constraint and a \$12 constraint should not have a similar effect. Again, this makes it less likely that the incorrect 4 balls against 4 weighing will be chosen in the \$8 condition, in turn resulting in more correct solutions. Therefore, the expectation is that, for the \$8 condition compared with the no constraint and \$12 conditions, because the incidence of criterion satisfying (4v4) moves should be lower, the incidence of correct solutions should be greater. In the 7-ball condition, unlike the 9-ball, the \$8 constraint does not cause immediate criterion failure, allowing the criterion satisfying 3v3 weighing with \$2 to spare. We predicted that the incidence of the criterion satisfying 3v3 move and the ultimate solution rates should be the same under \$8, no and \$12 cost constraints.

As indicated, Experiment 2 introduced a further change from Experiment 1, by adding a no cost constraint to the \$8 and \$12 cost constraints. While the difference in costs had some effect in the first experiment, it remains unclear whether the \$12 constraint had no effect or simply a lesser effect than the \$8 constraint. Introducing no cost constraint would clarify this issue.

3.1. Method

Participants. Study booklets were distributed to 165 adults attending a University open day presentation for prospective Psychology students (25 for each experimental condition). Once blank or incomplete booklets were discarded, data were collected from 156 adults (58 male, 98 female, mean age 22.0 years, range 17-54 years), with between 20 and 25 in each experimental condition.

Materials. The 7-ball and 9-ball source problems were identical to those used in Experiment 1, except for the addition of a variant for each problem in which the cost constraint information was removed. The ‘nuclear rods’ transfer problem is described in Appendix 1.

Design and Procedure. Participants were assigned to either 7-ball plus no cost constraint, 7-ball plus \$8, 7-ball plus \$12, 9-ball plus no constraint, 9-ball plus \$8 or 9-ball plus \$12 condition by alternating the order in which booklets were distributed across participants. In all other respects the design and procedure were identical to Experiment 1.

3.2. Results and Discussion

Source solution proportions. For the 7-ball source, solutions rates were 65% (13/20) under no cost constraint, 64% (14/22) under \$8, and 74% (17/23) under \$12. For the 9-ball condition, the corresponding rates were 24% (5/21), 56% (14/25), and 18% (4/22). A logistic regression using Source, Constraint, and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 133) = 25.01, p < .001$. Only the interaction between Source and Constraint (Wald = 6.03, $p = .049$) was a significant predictor in the model, with the \$8 constraint resulting in higher solution rates than no cost and \$12 in the 9-ball condition, but not the 7-ball. Constraint (Wald = 5.07, $p = .079$) and Source (Wald = .283, $p = .595$) were not significant predictors in the model.

Source first weighing frequencies. For the 7-ball source, the frequencies of criterion-satisfying first weighings were 75% (15/20) under no cost constraint, 73% (16/22) under the \$8 constraint and 83% (19/23) under the \$12. For the 9-ball source, the corresponding rates were 91% (19/21), 48% (12/25) and 95% (21/22). A logistic regression using Source, Constraint, and the interaction between these factors as

predictors yielded a significant model, $\chi^2(1, N = 133) = 18.94, p = .002$. Again, only the interaction between Source and Constraint (Wald = 6.14, $p = .038$) was a significant predictor in the model, with the \$8 constraint resulting in fewer criterion-satisfying first weighs than the \$12 and no cost constraints in the 9-ball condition, but not in the 7-ball. Constraint (Wald = 4.48, $p = .107$) and Source (Wald = 2.83, $p = .089$) were not significant predictors in the model.

Nuclear rods target problem. For those previously given the 7-ball source, solutions rates for the nuclear rods target problem were 20% (4/20) for the no cost constraint source, 27% (6/22) for \$8 and 22% (5/23) for \$12. For those given the 9-ball source, the corresponding rates were 38% (8/21), 72% (18/25) and 23% (5/22). A logistic regression using Source, Constraint, and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 162) = 18.74, p = .002$, with Source (Wald = 10.28, $p = .002$) and the interaction between Source and Constraint (Wald = 6.16, $p = .046$) significant predictors in the model. Constraint (Wald = 4.98, $p = .083$) was not significant in the model.

Target first weighing frequencies. The frequencies of criterion-satisfying (4v4) first moves for the nuclear rods target following the 7-ball source were 80% (16/20) under no source cost constraint, 77% (17/22) for the source \$8 constraint and 65% (15/23) under the source \$12. (It should be stressed that cost constraints were imposed in the source problems only, not in the target problem phase. For the 9-ball condition, the corresponding rates were 67% (14/21), 32% (8/25) and 73% (16/22). A logistic regression using Source, Constraint, and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 162) = 15.59, p = .008$, with Source (Wald = 8.85, $p = .008$) and the interaction between Source and Constraint (Wald = 6.30, $p = .043$) significant predictors in the model. Again, Constraint (Wald = 5.41, $p = .020$) was not significant in the model.

=.067) was not significant in the model.

Comparisons with control condition

Experiment 2 included a control condition whose participants received the nuclear rods problem without prior exposure to either the 7-ball or 9-ball problems or to the cost constraints. This allowed tests of whether the transfer effect observed above was positive, with prior exposure to the 9-ball condition improving subsequent performance on the 8-rods problem, or negative, with prior exposure to the 7-ball condition inhibiting performance on the 8-rods problem.

To make these comparisons, Chi squared analyses were conducted separately for the no cost constraint, \$8 and \$12 cost conditions, comparing solution rates to the 8-rods problem for 9-ball training, 7-ball training and no training (control) conditions. Because there was no cost constraint manipulation in the control condition, the same control data was used in each of these comparisons. For no cost constraint, no significant difference was found among the mean solution rates for 7-ball, 9-ball and no training, with means of 20% (5/20), 38% (8/21) and 14% (3/21), respectively, $\chi^2(1, N = 62) = 2.292, p = .318$. Similarly, for the \$12 constraint, no significant difference was found among the mean solution rates for 7-ball, 9-ball and no training conditions, with means of 22% (5/23), 23% (5/22) and 14% (3/21), respectively, $\chi^2(1, N = 66) = .960, p = .619$. For the \$8 constraint, however, there was a significant overall effect, $\chi^2(1, N = 71) = 22.65, p < .001$, with means of 27% (6/22), 72% (18/25) and 14% (3/21) for 7-ball, 9-ball and no training conditions. The findings support the conclusion that 9-ball \$8 training had positive transfer effects on subsequent performance on the 8-rods problem.

4. Experiment 3

In Experiments 1 and 2, evidence for spontaneous transfer was found, and this transfer seems mediated by experience of the source solution process. Experiment 1 found spontaneous transfer between superficially similar analogs, and Experiment 2 found it between conceptually similar but superficially different analogs. Experiment 3 explored the degree of conceptual and superficial distance still further, by using a problem variant for the transfer problem, the ‘sheepdog’ problem, which reduced further both relational and attributional similarities between source and analogue problems. In the sheepdog problem, the weigh scale was replaced by two barns in which sheep can be left overnight, the heavy ball was replaced by a ‘killer dog’ that had to be detected among 9 others (or 8 in the transfer problem), and the cost constraint was replaced by a ‘tablet’ that had to be given to dogs that were placed inside the barns. The sheepdog problem increases the analog distance from the n-balls problem still further than the nuclear rods problem by adding not only attributional differences but also an additional relational step: whereas the outcome of the test is revealed directly through effects on the primary attribute (balls/rods) in the latter problems, in the sheepdog problem the effect is revealed not on the primary attribute (the dogs) but on a secondary attribute (the status of the sheep in each barn).

A second issue examined in this experiment concerned the effect of problem difficulty on analogical transfer. A frequent finding in analogical transfer research is that stronger transfer effects are found if the source problem is more difficult than the target problem (e.g., Keane, 1997). In this experiment, the order of problems was varied systematically, to see if there were differences in source problem difficulty in training and, if so, whether this affected transfer performance. Therefore, approximately 50% of participants received the 9-ball problem as source and the

sheepdog problem as target, and 50% the reverse order. Here the prediction was of a main effect of cost constraint with higher success rates in the target problem and lower incidence of maximizing first weighs for those who had previously worked under the \$8 cost constraint, since the former will have had more direct exposure to criterion-failure than those trained under the \$0 and \$12 constraint.

As in Experiment 2, imposing an \$8 constraint in the 9-ball/9-sheepdogs conditions should cause immediate criterion failure, since the preferred 4 against 4 first testing will exhaust the available resources, whereas neither the \$0 nor \$12 constraint should have a similar effect. Again, this makes it less likely that the incorrect 4 balls/sheep against 4 testing will be chosen in the \$8 condition, in turn resulting in more correct solutions. Therefore, the expectation is that, for the \$8 condition compared with the \$0 and \$12 conditions, because the incidence of criterion satisfying (4v4) moves should be lower, the incidence of correct solutions should be greater. The experimental predictions were therefore for a significant main effect of cost constraint, with the \$8 constraint, but not the \$0 or \$12, having a facilitative effect on solution rate and an inhibitory effect on maximizing moves.

4.1. Method

Participants. One hundred and twenty first-year undergraduate students (42 male, 78 female, mean age 19.9 years, range 17-29 years) were tested as part of an introductory Psychology 101 class.

Materials. The 9-ball source problem was identical to that used in Experiment 2. A second source problem, the ‘sheepdog’ problem, shown in Appendix 1, was also used.

Design and Procedure. Participants were assigned to either 9-ball plus no cost constraint, 9-ball plus \$8 or 9-ball plus \$12, sheepdog plus no constraint, sheepdog

plus \$8, or sheepdog plus \$12 condition by alternating the order in which booklets were distributed across participants. In all other respects the design and procedure were identical to Experiments 1 and 2.

4.2. Results and Discussion

Source solution proportions. For the 9-ball training condition, solutions rates were 25% (5/20) under no constraint, 53% (10/19) under the \$8 constraint, and 15% (3/20) under the \$12 cost constraint. For the 9-sheepdogs training condition, the corresponding rates were 14% (3/21), 50% (10/20), and 25% (5/20). A logistic regression using Source (9-ball, sheepdogs), Constraint (no constraint, \$8/8 tablets and \$12/12 tablets), and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 120) = 20.41, p = .001$, with Constraint (Wald = 10.43, $p = .005$) the only significant predictor in the model. Source (Wald = .172, $p = .678$) and the interaction between Source and Constraint (Wald = .899, $p = .638$) were not significant.

Source first move frequencies. For the 9-ball condition, the frequencies of criterion-satisfying first weighings were 80% (16/20) under the \$0 constraint, 53% (10/19) under the \$8 constraint and 85% (17/20) under the \$12. For the 9-sheepdogs condition, the corresponding rates were 81% (17/21), 55% (11/20) and 75% (15/20). Although these frequencies are in the predicted direction, a logistic regression with Source, Constraint, and the interaction between these factors as predictors did not yield a significant model, $\chi^2(1, N = 120) = 9.38, p = .095$. However, a Chi squared test of the Constraint factor was significant, $\chi^2(1, N = 120) = 8.72, p = .013$, with the \$8/8 tablets constraint resulting in a lower incidence of 4v4 moves than no constraint and \$12/12 tablets.

Target solution proportions. For those given the 9-sheepdogs source, solution rates to

the 9-ball target were 19% (4/21) under the source no constraint, 80% (16/20) under \$8, and 30% (6/20) under \$12. For the 9-balls source, the corresponding solution rates to the sheepdogs target were 15% (3/20), 58% (11/19), and 15% (3/20). (Again, it should be noted that constraints were imposed in the source problem but not in the target.) A logistic regression using Source, Constraint, and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 120) = 32.00, p < .001$, with Constraint (Wald = 14.58, $p < .001$) the only significant predictor in the model. Source (Wald = 1.25, $p = .264$) and the interaction between Source and Constraint (Wald = .521, $p = .771$) were not significant.

Target first weighing frequencies. For the 9-sheepdogs source, the frequencies of criterion-satisfying first weighs on transfer to the 9-ball target were 81% (17/21) under no constraint, 30% (6/20) under \$8 and 63% (12/19) under \$12. For the 9-ball source, the corresponding rates on transfer to the sheepdogs target were 71% (12/17), 42% (8/19) and 67% (12/18). A logistic regression using Source, Constraint, and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 120) = 17.10, p = .004$, with Constraint (Wald = 10.24, $p = .006$) the only significant predictor in the model, with the \$8/8 tablets constraint resulting in a lower incidence of 4v4 moves than no constraint and \$12/12 tablets. Source (Wald = .114, $p = .736$) and the interaction between Source and Constraint (Wald = .758, $p = .685$) were not significant.

In Experiment 3, the analog problem required identifying a rogue sheepdog. In terms of analogical mappings to the n -ball problem, dogs mapped to balls, barns and leashes played the role of the balance, and the presence/absence of a dead sheep in a barn was the analogical equivalent of a difference in weight. While these structural mappings were similar to those of the problems used in the previous

experiments, there were also some differences. The objects ‘pans’ and ‘bins’ of Experiments 1 and 2 had attributes ‘on/off’ or ‘up/down’ that led directly to the test result. In contrast, the sheepdog problem had the object ‘barn’ with the attribute ‘contains sheep’; the sheep represented an added problem object with the attribute ‘alive/dead’. As such, the sheepdog problem is structurally similar to, but not identical with, the target problem, since test result must be inferred from an additional level of object-attribute relations. On this basis, the problem is a more distant analog of the n -ball problem than is the nuclear rod problem of Experiment 2, and analogical transfer increased from relatively near in Experiment 1 to relatively far in Experiment 3. Nevertheless, the results indicated that spontaneous transfer did occur in Experiment 3.

While the two source problems for Experiment 3 both had nine balls (or sheepdogs) and only the cost constraint varied, results were similar to those of the previous experiments. Solution rates and first moves for source problems followed a similar pattern, in that solution rates were significantly higher and maximizing first moves were significantly lower under the tight constraint than under the lax or no constraint conditions. A similar pattern emerged on transfer, with higher solution rates and lower maximizing first moves for those trained under the tight constraint.

5. Experiment 4

The first three experiments found evidence of spontaneous transfer from n -ball problems to apparently more distant analogs across successive experiments, with transfer to a different n -ball problem in Experiment 1, to a problem involving nuclear rods in Experiment 2, and to a pack of sheepdogs with one rogue in Experiment 3.

In all three experiments, a common element linking source and target problem solutions was that they required avoiding a maximizing move, and one interpretation of the findings is that the observed transfer effects resulted from participants learning not to follow the impulse to maximize. That said, the results are open to other interpretations, the strongest candidate being that transfer was facilitated by structural alignment of source and target problems due to their common underlying structural design. Other possible explanations include the transfer of problem-specific strategies, such as to have no more than three objects for the second move. Similarly, differences in transfer to the n-ball task could have resulted from learning different specific strategies in the training conditions, such as “have no more than 3 balls on each side” in the 9-ball condition versus “divide the balls as evenly as possible” in the 7-ball. It would be possible to generate problem- and condition- specific explanations across all of the experiments, and while arguably less parsimonious than our preferred account, such explanations may nevertheless be correct. None of these alternatives to the transfer of non-maximization hypothesis offers an immediate explanation for the effects of tight versus lax constraints. Yet, the possibility remains that manipulating this constraint in some way interacts with the utility of solution-based knowledge or structural alignment.

To seek more direct evidence for the transfer of non-maximization hypothesis, Experiment 4 used source and target tasks that had no apparent structural overlap, thereby removing the basis for a structural alignment explanation. In addition, the tasks had few if any obvious superficial similarities that might be used to generate problem-specific strategies.

Unlike the previous experiments, Experiment 4 exposed participants to two training problems, the 9-ball and the 9-Sheepdog problems. The intention here was to

amplify any learning effects by supplying more than one training experience (Gick & Holyoak, 1983). In addition, both source tasks included the cost constraints used previously, to discourage maximization under a tight but not a lax constraint. As described in the Introduction, the target task was the Cheap Necklace Problem (CNP), which involved different superficial elements and a different structural design than either source problem.

5.1. Method

Participants. Two hundred and ten second-year undergraduate students (61 male, 149 female, mean age 20.6 years, range 19-30 years) were tested as part of a Psychology Research Methods class.

Materials. The 9-ball and 9-sheepdogs source problems were identical to those used in Experiment 3. Duncker's (1945) Radiation problem and its analog the Fortress problem (Gick & Holyoak, 1983) were used as alternative source analogs for control conditions. The Cheap Necklace Problem (CNP) was used as the target problem in all conditions. These problems are shown in Appendix 2.

Design. Participants in experimental conditions were assigned to either balls-first or sheepdogs-first conditions, participants in the former receiving the 9-balls problem as the first source problem and the 9-sheepdogs problem as the second source problem, participants in the latter receiving the source problems in the opposite order. Participants in each of these Problem Order conditions were further assigned to either lax or tight Constraint conditions (i.e., 9-ball with either \$12 or \$8; 9-sheepdog with either 11 or 9 tablets).

In order to assess the impact on CNP solution rates in the absence of training with ball-variant source analogues, three control conditions were included. In the first, to allow comparison of experimental conditions against a no-training baseline,

participants solved CNP alone. In a second control, participants solved the Radiation problem followed by the CNP. In a third control, participants solved the Fortress problem, followed by the Radiation problem, followed by the CNP.

Procedure. Participants were tested in groups of approximately 30, working individually during three-hour laboratory classes held as part of a Research Methods course. Experimental and control conditions were assigned to different laboratory classes. Participants were assigned to conditions by alternating the order in which problem booklets were distributed. Participants attempted the first problem at the start of the laboratory class, the second problem after one hour, and the third problem after two hours. Participants attempting only one or two problems did so in the second and/or third testing times. In all other respects the procedure was identical to Experiments 1, 2 and 3.

5.2. Results and Discussion

First training problem solution proportions. For the 9-ball-first condition, solutions rates were 49% (19/39) under the tight (\$8) constraint, and 18% (7/40) under the lax (\$12) constraint. For the 9-sheepdogs training condition, the corresponding rates were 46% (18/39) under the tight constraint, and 17% (6/36) under the lax constraint. A logistic regression using Source (9-ball, 9-sheepdogs), Constraint (no constraint, \$8/9 tablets and \$12/11 tablets), and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 154) = 16.73, p = .001$, with Constraint (Wald = 8.15, $p = .004$) a significant predictor in the model. Source (Wald = .051, $p = .821$) and the interaction between Source and Constraint (Wald = .003, $p = .954$) were not significant.

First training problem, first move frequencies. For the 9-ball condition, the frequencies of maximising first weighs were 62% (24/39) under the tight constraint

and 90% (36/40) under the lax constraint. For the 9-sheepdogs condition, the corresponding rates were 49% (19/49) and 88% (32/36). A logistic regression with Source, Constraint, and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 154) = 25.27, p < .001$, with Constraint (Wald = 7.73, $p = .005$) a significant predictor in the model. Source (Wald = 1.29, $p = .256$) and the interaction between Source and Constraint (Wald = .211, $p = .646$) were not significant.

Second training problem solution proportions. For the 9-ball-first condition, solutions rates were 51% (20/39) under the tight constraint, and 33% (13/40) under the lax constraint. For the 9-sheepdogs-first training condition, the corresponding rates were 52% (20/39) under the tight constraint, and 33% (12/36) under the lax constraint. A logistic regression with Source, Constraint, and the interaction between these factors as predictors did not yield a significant model, $\chi^2(1, N = 154) = 5.39, p = .146$.

Second training problem, first move frequencies. For the 9-ball-first condition, the frequencies of maximising first weighs were 44% (17/39) under the tight constraint and 70% (28/40) under the lax constraint. For the 9-sheepdogs-first condition, the corresponding rates were 44% (20/39) and 69% (29/36). A logistic regression with Source, Constraint, and the interaction between these factors as predictors did not yield a significant model, $\chi^2(1, N = 154) = 2.698, p = .441$.

Target solution proportions. For those given the 9-sheepdogs source first and the 9-ball second, solution rates to the CNP were 44% (17/39) under the tight constraint, and 25% (9/36) under the lax constraint. For those given the 9-ball source first and the 9-sheepdogs second, solution rates to the CNP were 49% (19/39) under the tight constraint, and 20% (8/40) under the lax constraint. A logistic regression using

Source, Constraint, and the interaction between these factors as predictors yielded a significant model, $\chi^2(1, N = 154) = 10.293, p < .016$, with Constraint (Wald = 6.884, $p = .009$) the only significant predictor in the model. Source (Wald = .206, $p = .650$) and the interaction between Source and Constraint (Wald = .478, $p = .489$) were not significant.

Comparisons with control conditions. Solution frequencies for participants receiving the CNP alone were 25% (5/20). For participants who received the Radiation problem as a source problem, solution rates for the CNP were 16% (3/19). For participants who received both the Fortress and Radiation problems as source problems, solution rates for the CNP were 18% (3/17). A Chi squared analysis showed no significant difference in solution rates between these conditions, $\chi^2(1, N = 56) = 0.585, p = .746$.

To compare against experimental conditions, the control conditions were collapsed into a single control. Similarly, since there were no significant effects in any analysis involving Problem Order, CNP solution frequencies were summed across Problem Order conditions. A Chi squared analysis of CNP solution rates between control (11/56) and experimental conditions with tight constraint source problems (36/78) was significant, $\chi^2(1, N = 134) = 10.061, p = .002$. A Chi squared analysis of CNP solution rates between control (11/56) and experimental conditions with lax constraint source problems (17/76) was not significant, $\chi^2(1, N = 143) = 0.371$.

6. General Discussion

In this paper we tested a theory of spontaneous analogy based on transfer of experience of heuristic change. The theory extends an approach to understanding insight problem solving (CSPT) developed to account for performance on a range of insight problems that includes the 9-dot and related problems (MacGregor et al, 2001)

and a variety of coin problems (Ormerod et al, 2002; Chronicle et al, 2004). More recently, we extended the theory to the n -ball problem (Ormerod et al, 2013), where it was successful in predicting problem difficulty and the proportions and sequence of first moves in a multi-trial experiment. Through the construct of zero lookahead, the theory also explained the “irrational” moves previously reported by Simmel (1953).

The present article extends this work by using a transfer paradigm to test an account of the conditions under which spontaneous analogy might arise. Within experiments, participants received additional constraints that were irrelevant to the correct problem solution path but that systematically altered the experience of applying problem heuristics. The constraints varied the apparent availability of resources to solve source problems, with exact amounts that discouraged a maximization heuristic or excess amounts that allowed free choice of heuristic. Across the experiments, materials were varied to increase the distance between source and target problems, to test the range of spontaneous analogical transfer with this type of problem.

The first experiment manipulated source problem (7-ball versus 9-ball) and cost constraint (exact versus excess) and tested transfer effects using the 8-ball problem. The source results indicated that the manipulations had an interactive effect, with the 9-ball tight-cost constraint condition having the highest solution rates, as predicted. The proposed mechanism for this effect, a reduction in maximizing first weighs, was also supported, with the 9-ball tight-cost constraint condition having the lowest frequency of maximizing weighs. The effects carried over to the transfer task, where solution rates for the 8-ball problem were highest for those who had received the 9-ball tight-cost constraint as the source problem. For first weighs in the transfer task, frequencies of maximizing weighs were also in the predicted direction (with the

9-ball and tight-cost condition showing fewer maximizing weighs than the other conditions).

In Experiment 2, the target problem involved testing 8 nuclear rods to isolate one creating excess heat, with plutonium rods in place of balls, a testing device instead of a balance, and a difference in heat instead of weight. Although structurally identical, source and target problems had more surface differences than in the first experiment. In addition, the experiment included a no-cost condition and employed a control condition in which participants attempted the nuclear rods problem with no prior training. This allowed us to examine whether any transfer effects that occurred were positive or negative. Transfer performance on the nuclear rods problem indicated that solution rates were highest for the 9-ball tight (\$8) constraint condition, with fewer maximizing first weighs under 9-ball training than 7-ball training under tight cost rather than the lax or no cost constraints. Performance under the 9-ball tight constraint was the only condition significantly better than the control in which participants received only the 8-nuclear rods problem.

In Experiment 3, the analog problem involved isolating a sheepdog that had turned sheep killer, where dogs substituted for balls, barns and leashes substituted for the balance, and the presence or absence of a dead sheep in a barn substituted for a difference in weight. The objects ‘pans’ and ‘bins’ of Experiments 1 and 2 had attributes ‘on/off’ or ‘up/down’ that directly yielded the test result. In contrast, the sheepdog problem had an object ‘barn’ with the attribute ‘contains sheep’; the sheep presented an additional problem object with the attribute ‘alive/dead’. As such, the sheepdog source problem is structurally similar to, but not identical with, the target problem, since test result must be inferred from an additional level of object-attribute relations. On this basis, the problem is a more distant analog of the n -ball problem

than is the nuclear rod problem of Experiment 2, and analogical transfer increased from relatively near in Experiment 1 to relatively far in Experiment 3. Nevertheless, spontaneous transfer did occur in Experiment 3: higher solution rates and lower maximizing first moves were achieved by those trained under the tight constraint.

It remains possible that the results of the first three experiments were due to factors other than the transfer of heuristic change. Specifically, structural alignment between source and target problems may have played a determining or contributing role. Experiment 4 used a transfer task that engages a maximizing heuristic but otherwise appeared to bear little or no resemblance to the training tasks. The results were consistent with those of the previous results in the effects of tight and lax source problem constraints on the training performance and on the transfer task: higher transfer solution rates were observed under tight constraint training than under lax. In addition, solution rates for the transfer task were significantly higher than control solution rates for tight constraint training, but not lax.

As noted in the introduction, spontaneous analogical transfer is relatively rare, the main impediment being not the lack of structural fit between base and target problems, but the failure of problem solvers to recognize the analogy (Day & Goldstone, 2011). This recognition failure is particularly likely if surface features are dissimilar, which makes it all the more surprising that transfer was observed in Experiments 2, 3 and 4, which involved apparently high levels of surface dissimilarity. Also as noted above, transfer might be mediated by activation and application of previously acquired skills and knowledge (e.g., Gick and Holyoak, 1983) or through inhibiting a response (e.g., Detterman, 1993). The current results are consistent with the latter, suggesting that the transfer observed also involved a change in problem-solving heuristic, not simply mapping problem structures or elements.

The 9-balls source and Cheap Necklace target problems of Experiment 4 shared an attributional alignment, in that both contained a cost constraint. However, the cost attributes served different functions in each problem, and did not align relationally. Thus, both a Thorndyikian account based on the principle of common elements and a structural alignment account might predict a null or negative effect of transfer, since attempts to map solution components from source to target based on attributional mappings would necessarily fail. Moreover, while the Radiation and Fortress control training problems did not involve cost constraints, they did share a potentially important relational alignment with the target problem, in that the solutions to the Radiation, Fortress, and Cheap Necklace problems all required dividing a larger entity into smaller components. It is unclear why this relational similarity failed to trigger solutions to the CNP while inhibition of maximization did.

A question remains as to whether the inhibition of maximization observed in the present experiments represents a separate phenomenon or whether it should be regarded as an example of constraint relaxation. It could be argued that maximization is a form of hill-climbing under the constraint of making maximum progress with each move, and that the inhibition of maximization is equivalent to the relaxation of this constraint to maximize. Such an interpretation has the advantage of forging a further rapprochement between the CSPT approach to problem solving and Representational Change Theory (RCT), which proposes constraint relaxation as one general mechanism underlying problem solving and transfer (Knoblich, Ohlsson, Haider, & Rhenius, 1999). While, in the interests of greater generality, it is tempting to place inhibition of maximization in the same category as constraint relaxation, we see at least two problems with doing so, which we elaborate below.

The first problem is related to the fact that some constraints are more difficult

to relax than others (Knoblich et al, 1999; Knoblich, Ohlsson & Raney, 2001). RCT accounts for this by proposing that the probability of a constraint being relaxed depends on its *scope*, defined in terms of the breadth of problem representation over which a constraint applies. Specifically, constraints that have a very general application will be more difficult to relax than constraints that have only a local application. The problem arising from this is that maximization appears to be a highly general ‘weak-method’ heuristic (Newell & Simon, 1972; Anderson, 2014), belonging to the category of search heuristics that includes hill-climbing, gap reduction and means-ends analysis. As such, it would appear to have a wide scope, and should be correspondingly difficult to relax. From an adaptation perspective it makes sense that a widely applicable constraint would be difficult to relax. However, the present findings demonstrate inhibition of maximization, suggesting either that that inhibition of maximization is *not* constraint relaxation, or constraint relaxation does *not* operate as RCT proposes.

A second problem arises because of RCT’s assumption that “...once relaxed, constraints stay relaxed” (Knoblich et al, 1999, p. 1539). Given the nature of maximization as a general heuristic for search space reduction, inhibiting it indefinitely would be of questionable functional value. In contrast to a binary on-off state, the approach we have taken proposes that maximization is reduced gradually, through the operation of the minimization heuristic. In the first three experiments, problem solution required suppressing maximization to a relatively small degree in that instead of the maximizing 4 v 4 weighing (or equivalent test) target solution required 3 v 3 weighs, the next-to-maximum balanced alternative. It may be that for problems requiring a much greater reduction in maximization transfer effects would be more difficult to obtain, would have a more restricted range (i.e. near transfer but

not far) and would persist for a shorter duration.

Nevertheless, the transfer effects in the present experiment were observed over a relatively wide range of transfer distance and following an activity-filled period of 45 minutes. The question remains open of how far the effect generalizes and for how long it persists.

The suppression of a maximizing heuristic appears to yield benefits for spontaneous analogy. Yet, anything other than a temporary suppression of a maximization heuristic would likely have deleterious effects on problem-solving in the longer term. It may be that analogical mapping of structural and superficial properties between source and target problems plays an important role in determining the ‘scope’ for suppressing maximization in the longer term: if suppression helps solve a source, a similar approach may be applied to a target if and only if structural and/or superficial mappings can be made between the problems. Further research is needed to explore a potentially symbiotic relationship between analogical mapping and scope of heuristic change.

The results of the current research suggest that factors beyond problem structures, relations and elements play an influential role in analogical problem solving. If correct, this could have important implications for efforts to apply analogical learning in education and training. The role of analogy in these domains is perhaps less straightforward than previously assumed. For example, Kalyuga, Chandler, Tuovinen, & Sweller (2001) showed that, with increasing domain expertise, guided (rather than spontaneous) analogy through worked examples decreases in effectiveness compared with weak-method problem-solving heuristics. More recently, Richey, Zepeda & Nokes-Malach (2015) have shown that self-explanation can lead to higher levels of transfer than explicit provision of analogical mapping materials.

These counter-intuitive findings, combined with our results concerning the heuristic mediation of spontaneous analogy, suggests that efforts to apply psychological theories of analogy to education and training must await a fully interactive model of the heuristic and knowledge-based factors underlying analogical problem-solving. A move towards such a theory might also enrich the development of neurologically plausible models of analogy, adding more strategic components (e.g., Barutta, Guex, & Ibáñez, 2010) to current models that focus on the role of pre-frontal cortex in identifying structural representations during analogy (e.g., Speed, 2010).

7. References

- Anderson, J. R. (2014). *Rules of the mind*. New York: Psychology Press.
- Anderson, J. R., Farrell, R., & Sauers, R. (1984). Learning to program in LISP. *Cognitive Science*, 8, 87–129.
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128, 612-637.
- Barutta, J., Guex, R., & Ibáñez, A. (2010). Does the PFC model of analogy account for decision making, problem solving, reasoning, flexibility, adaptability, and even creativity? *Cognitive Neuroscience*, 1, 142-143.
- Bassok, M. (1990). Transfer of domain-specific problem-solving procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 522-533.
- Bearman, C.R., Ball, L.J., & Ormerod, T.C. (2007). The structure and function of spontaneous analogising in domain-based problem solving. *Thinking and Reasoning*, 13, 273-294.
- Bearman, C., Ormerod, T.C., Ball, L.J., & Deptula, D. (2011). Explaining away the negative effects of evaluation on analogical transfer: The perils of premature evaluation. *Quarterly Journal of Experimental Psychology*, 64, 942-959.
- Blanchette, I., & Dunbar, K. (2000). How analogies are generated: The roles of structural and superficial similarity. *Memory & Cognition*, 28, 108-124.
- Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1147-1156.
- Catrambone, R., Craig, D. L., & Nersessian, N. J. (2006). The role of perceptually represented structure in analogical problem solving. *Memory & Cognition*, 34,

1126-1132.

- Clement, J.J. (1987). Generation of spontaneous analogies by students solving science problems. Chapter 23 in D. M. Topping, D.C. Crowell, & V.N. Kobayashi (Eds.). *Thinking across cultures*, pp.303-308.
- Chu, Y., Dewald, A., & Chronicle, E.P. (2007). Theory-driven hints in the Cheap Necklace Problem: A preliminary investigation. *Journal of Problem Solving*, 1, 18-32.
- Day, S. B., & Goldstone, R. L. (2011). Analogical transfer from a simulated physical system. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 551-567.
- Day, S.B., & Goldstone, R.L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist*, 47, 153-176.
- Detterman, D. (1993). The case for the prosecution: Transfer as an epiphenomenon. In D. Detterman & R. Sternberg (Eds.), *Transfer on trial: Intelligence, Cognition, and Instruction* (pp. 1-24). Norwood, NJ: Ablex.
- Didierjean, A., & Nogry, S. (2004). Reducing structural-element salience on a source problem produces later success in analogical transfer: What role does source difficulty play? *Memory & Cognition*, 32, 1053-1064.
- Duncker, K. (1945). On problem solving. *Psychological Monographs*, 58, 1-113.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.
- Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52, 45-56.
- Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in

- transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, 25, 524–575.
- Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, 10, 277-300.
- Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, 12, 306-355.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1–38.
- Holyoak, K. J. (1985). The pragmatics of analogical transfer. *The Psychology of Learning and Motivation*, 19, 59-87.
- Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, 15, 332-340.
- Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13, 295-355.
- Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology*, 93, 579-588.
- Keane, M.T. (1997). What makes an analogy difficult ? The effects of order and causal structure in analogical mapping. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 23, 946-967.
- Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 25, 1534-1555.
- Knoblich, G., Ohlsson, S., & Raney, G.E. (2001). An eye movement study of insight problem solving. *Memory and Cognition*, 29, 1000–1009.

- MacGregor, J. N., & Cunningham, J. B. (2008). Rebus puzzles as insight problems. *Behavior Research Methods*, 40, 263-268.
- MacGregor, J.N., Ormerod, T.C., & Chronicle, E. (1999). Spatial and contextual factors in human performance on the travelling salesperson problem. *Perception*. 28, 1417-1427.
- Medin, D. L., & Ross, B. H. (1989). The specific character of abstract thought: Categorization, problem solving, and induction. Chapter in R.J. Sternberg (Ed), (1989). *Advances in the Psychology of Human Intelligence*, Vol. 5., (pp. 189-223). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc, 236 pp.
- Monaghan, P., Sio U-N, Lau S.W., Woo, H.K., Linkenauger, S.K., & Ormerod, T.C. (2015). Sleep promotes analogical transfer in problem solving. *Cognition*, 143, 25-30. <http://dx.doi.org/10.1016/j.cognition.2015.06.005>.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs: Prentice-Hall.
- Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 398-415.
- Ormerod, T. C., MacGregor, J. N. & Chronicle, E. P. (2002). Dynamics and Constraints in Insight Problem Solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 791-799.
- Ormerod, T.C., MacGregor, J.N., Chronicle, E.P., Dewald, A.D., & Chu, Y. (2013). Act first, think later: Effects of maximization, minimization and lookahead on inferential planning in problem-solving. *Memory & Cognition*, 41, 1096-1108. doi: 10.3758/s13421-013-0318-5.
- Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to

- categorize statistics word problems. *Journal of Educational Psychology*, 88, 144–161.
- Reed, S. K. (1989). Constraints on the abstraction of solutions. *Journal of Educational Psychology*, 81, 532-540.
- Reed, S. K., Dempster, A., & Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 106-115.
- Reeves, L., & Weisberg, R. W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin*, 115, 381- 400.
- Richey, J. E., Zepeda, C. D., & Nokes-Malach, T. J. (2015). Transfer effects of prompted and self-reported analogical comparison and self-explanation. In D. Noelle and R. Dale (Eds.) *Proceedings of the 37th Annual Conference of Cognitive Science Society*, Austin, Texas: Cognitive Science Society.
- Ross, B. H. (1984). Reminders and their effects in learning a cognitive skill. *Cognitive Psychology*, 16, 371-416.
- Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 629-639.
- Russell, S. J., & Norvig, P. (2004). *Artificial Intelligence: A Modern Approach*. NJ: Prentice Hall.
- Schunn, C. D., & Dunbar, K. (1996). Priming, analogy, and awareness in complex reasoning. *Memory & Cognition*, 24, 271-284.
- Smith, S. M. (1995). Fixation, incubation, and insight in memory and creative thinking. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.). *The Creative Cognition Approach* (pp. 135-146). Cambridge, Mass.: MIT Press.

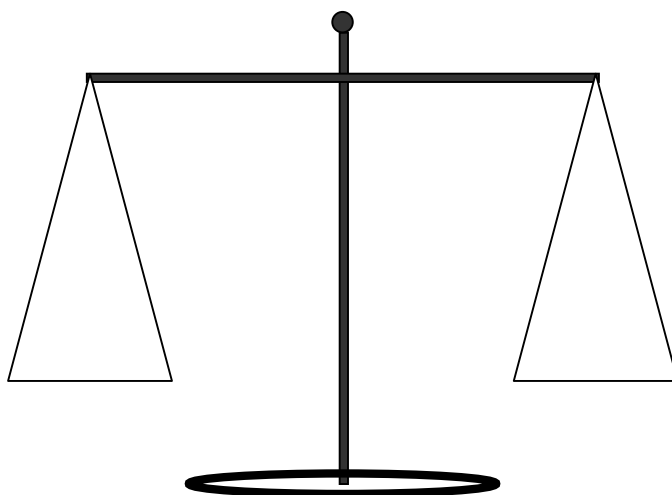
- Smith, S. M. & Blankenship, S. E. (1989). Incubation effects. *Bulletin of the Psychonomic Society*, 27, 311-314.
- Speed, A. (2010). Abstract relational categories, graded persistence, and prefrontal cortical representation. *Cognitive Neuroscience*, 1, 126–152.
- Thorndike, E. L. (1906). *Principles of teaching*. New York: Seiler.
- Weisberg, R. W., & Alba, J. W. (1981). An examination of the alleged role of "fixation" in the solution of several "insight" problems. *Journal of Experimental Psychology: General*, 110, 169-192.

Appendix 1. N-ball variant source and target problems

The n-balls problem (all experiments). The penultimate sentence of the first paragraph was omitted from zero cost conditions.

You have [seven/eight/ nine] balls that look identical, yet one of them is slightly heavier than the others (but the difference is too small to detect just by picking them up). You have a balance-scale, and you can use it only twice (that is to say, you have two weighings at your disposal). Also, it costs \$1 to weigh each ball, and you have a maximum of [\$8/\$12] to spend. Your task is to find out which one is heavier, using only two weighings and without spending more than [\$8/\$12].

1. In the space below, record the very first thing you thought of, for your first weighing. Draw in how many balls you want to put in the left balance pan, and how many in the right. It doesn't matter whether you feel you're right or wrong – just be honest about what you thought. If you want to provide an explanation or reason for your decision, use the “Comments” space below.



Comments:

2. Next, please record in the space below your best attempt at a solution, in whatever

way makes sense to you. That is, what two weighings would you make, and why?

On completion, all participants were shown the following solution information:

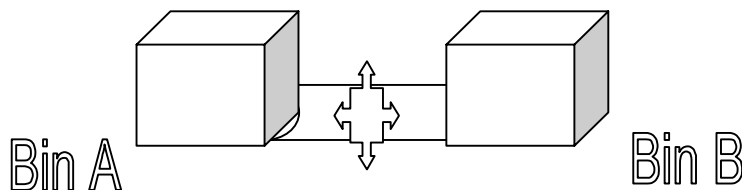
The solution to the problem you have just solved is as follows:

Put three balls on each side of the scales. If the scales balance, then the unweighed ball is the heavier one. Otherwise, it is one of the three balls that were on the downward side of the scale. Take two of the three balls that include the heavier one, and weigh one against the other. If the scales balance, then the unweighed ball is the heavier one. Otherwise, the ball on the side of the scales that dips downwards is the heavier one.

The Nuclear Rods problem (Target problem, Experiment 2)

“A nuclear reactor is in danger of exploding. There are eight plutonium rods in the reactor core, and one of them has a fault, invisible to the human eye, causing excess heat generation. There is a device that can test for the fault. The device has two ‘bins’ (A and B) into which rods are placed. To operate, you load a number of rods into each bin, and the device measures differences in heat production between the two bins. Unfortunately, you only have time for two tests before the reactor turns critical.

1. In the space below, record the very first thing you thought of, for your first test. Draw in how many rods you want to put in Bin A, and how many in Bin B. It doesn’t matter whether you feel you’re right or wrong – just be honest about what you thought. If you want to provide an explanation or reason for your decision, use the “Comments” space below.



Comments:

2. Next, please record in the space below your best attempt at a solution, in whatever way makes sense to you. That is, what two tests would you make, and why?”

The Sheepdog problem (Source/Target problem, Experiment 3. Paragraph 3 was omitted for zero cost-constraint conditions):

“A farmer has 8 dogs, which he uses to protect his flock of more than 1000 sheep from attack by wolves. Unfortunately, one dog has turned bad and has become a sheep killer. It will kill one sheep (exactly one, no more, no fewer) each night it is left with the flock. The other dogs won’t stop this from happening, or reveal the culprit. Now, the farmer expects to lose a few sheep from time to time, but he cannot afford to lose a sheep every single night of the year. The sheep killer is cunning and covers its traces by morning. It would never strike while the farmer is present, but the farmer cannot work all day and stay with the dogs all night, and there is no-one else who can keep watch.

The farmer has two barns in which he keeps his sheep to protect them from wolves in the winter. The dogs are placed in the barns with the sheep or can be tethered outside. It is now spring, and in two nights time the whole flock must be let out to spring pastures, since the farmer can’t afford to feed them in the barns any longer. They will need to be accompanied by the dogs for protection. If the killer dog isn’t found, it will eventually kill the entire flock and the farmer will starve.

Disease control is a key issue for the farmer, since the sheep are known to carry a disease that is deadly to dogs. To prevent the passage of disease from sheep to dogs, each night a dog is placed in the barn, the dog must be given a special tablet that will prevent it catching a disease from the sheep. Unfortunately, the farmer only has [eight/twelve] tablets left and cannot get hold of any more.

The farmer has no technology to help him identify the rogue dog (no

cameras, videos, mobile phones - nothing like that) - he is a simple peasant farmer. Nor can he get anyone else to help him, and he cannot stay with the dogs all the time. All he has is the sheep, the barns, the dogs, and short leads to tie up each dog, and he only has two nights to find the killer dog. How can he find out which dog is the killer before he has to let the sheep out to pasture in two nights time?"

Appendix 2. Alternative source analogs and target problems used in Experiment 4

The Fortress problem

A small country fell under the iron rule of a dictator. The dictator ruled the country from a strong fortress. The fortress was situated in the middle of the country, surrounded by farms and villages. Many roads radiated outward from the fortress like spokes on a wheel. A great general raised a large army at the border, vowing to capture the fortress and free the country from the dictator. The general knew that, if his entire army could attack the fortress at once, it could be captured. His troops were poised at the head of one of the roads leading to the fortress, ready to attack.

However, a spy brought the general a disturbing report. The ruthless dictator had planted mines on each of the roads leading to the fortress. The mines were set so that small bodies of men could pass over them safely, since the dictator needed to be able to move troops and workers to and from the fortress. However, any large force would detonate the mines. Not only would this blow up the road and render it impassable, but the dictator would then destroy many villages in retaliation. The good general knew that it takes at least a week to clear mines from a road, and he had only two days to launch the attack before the dictator would become aware of the threat and retaliate. A full-scale direct attack on the fortress appeared impossible.

What would you advise the good general to do in order to capture the fortress?

Write or draw your solution below:

Solution to the problem shown on the previous page is as follows:

The general divided his armies into small groups and dispatched each group to the head of a different road. When all was ready, he gave the signal and each group marched down a different road. Each group continued down its road so that the entire army arrived together at the fortress at the same time. In this way, the general captured the fortress and overthrew the dictator.

Duncker's (1945) radiation problem

Suppose you are a doctor faced with a patient who has a malignant tumour in his stomach. It is impossible to operate on the patient; but unless the tumour is destroyed the patient will die. There is a kind of ray that can be used to destroy the tumour. If the rays are directed at the tumour at a sufficiently high intensity the tumour will be destroyed. Unfortunately, at this intensity the healthy tissue that the rays pass through on the way to the tumour will also be destroyed. At lower intensities the rays are harmless to the healthy tissue but they will not affect the tumour either. What type of procedure might be used to destroy the tumour with the rays, and at the same time avoid destroying the healthy tissue?

Write or draw your solution below:

A solution to the problem shown on the previous page is as follows:

Place a number of different ray-emitting machines in different positions around the patient. Then simultaneously send low-intensity beams from each of the machines, so the rays all converge on the site of the tumour. The converging rays then have a sufficient combined intensity to destroy the tumour without damaging the surrounding tissue.

The cheap necklace problem

You have been given a necklace, but unfortunately it has broken into four pieces, as shown below (given state). You want to get it fixed (as shown in the goal state) so that it is a circular necklace again. A Jeweller offers to fix it, but charges £1 to break a link and £2 to make a link. You only have £9. Is it possible to mend the necklace for this amount of money, and if so how?

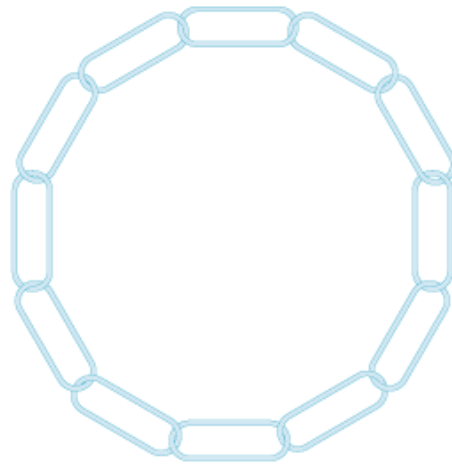
Chain A 

Chain B 

Chain C 

Chain D 

Given state

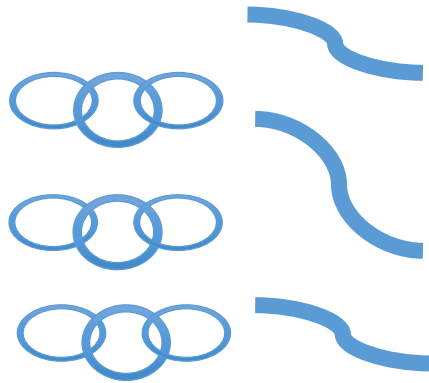


Goal state

Write or draw your solution here:

Solution to the cheap necklace problem:

Take one of the three-link lengths and break each link thus:



This will cost £3.

You can then join each chain length together to form a circle by using the broken ends as links that each cost £2. Thus the grand total is £9.